

Research Article

A 178 bp diagnostic D-loop fragment (OvisCR) to discriminate domestic sheep (*Ovis aries*) and European mouflon (*Ovis musimon*) in diet metabarcoding studies

Agathe Pirog^{1,2}, Juliette Lavarec¹, Sabine Rousselot¹, Cécile Kaerle¹, Christophe Duchamp³, Guillaume Queney¹

¹ ANTAGENE, Animal Genomics Laboratory, 6 allée du Levant, 69890 La Tour de Salvagny (Lyon), France

² French Agency for Biodiversity, Department of Research and Expertise, Parc d'affaires La Rivière, 8 Boulevard Albert Einstein, CS 42355 44323 Nantes, France

³ French Agency for Biodiversity, Department of Research and Expertise, Parc Micropolis, F-05000 Gap, France

Corresponding author: Agathe Pirog (paper@antagene.com)

Abstract

Accurate species identification is essential in metabarcoding diet studies, notably to distinguish domestic from wild species, as they have markedly different ecological and socio-economic implications. This is particularly true for the genus *Ovis*, where domestic sheep (*Ovis aries*) and European mouflon (*O. musimon*) often co-exist, but are indistinguishable with common mitochondrial markers due to insufficient variation. Misidentification of these taxa can skew carnivore predation assessments, notably for species like the grey wolf (*Canis lupus*), impacting wildlife conservation and management efforts.

Here, we developed a diagnostic D-loop fragment that effectively differentiates *O. aries* and *O. musimon* haplotypes in diet metabarcoding studies. Using reference samples from 46 European mouflons and 46 domestic sheep collected in the French Alps, we identified a 178 bp fragment (OvisCR) containing 43 SNPs, distinguishing four mouflon haplotypes and 35 sheep haplotypes. No haplotypes were shared between the two species, though one *O. musimon* haplotype was closely related to two *O. aries* haplotypes.

The performance of the OvisCR fragment was tested using 115 wolf scat samples. We sequenced a general metabarcoding marker (a 60–140 bp amplicon of the 12S region) alongside the new amplicon on the Ion Torrent platform to assess the ability of the latter to refine *Ovis* identification. Following sequencing, two bioinformatic pipelines (DADA2 and OBITools) were employed to compare their effectiveness in resolving *Ovis* operational taxonomic units (OTUs) or amplicon sequence variants (ASVs).

The two pipelines demonstrated similar accuracy, with the OvisCR fragment achieving *Ovis* species-level assignment for more than 81% of scats with *Ovis* reads detected, regardless of the pipeline used.

This new marker provides a robust and practical tool for metabarcoding applications, improving ecological inference in carnivore trophic studies and supporting biodiversity conservation management strategies. While effective in the French context, its broader geographic applicability requires prior validation using region-specific reference datasets that encompass the diversity of domestic sheep, European mouflon and other *Ovis* species.

Key words: *Canis lupus*, diet metabarcoding, mitochondrial control region, *Ovis aries*, *Ovis musimon*, species discrimination



Academic editor: Emre Keskin

Received: 18 September 2025

Accepted: 26 January 2026

Published: 11 March 2026

Citation: Pirog A, Lavarec J, Rousselot S, Kaerle C, Duchamp C, Queney G (2026) A 178 bp diagnostic D-loop fragment (OvisCR) to discriminate domestic sheep (*Ovis aries*) and European mouflon (*Ovis musimon*) in diet metabarcoding studies. Metabarcoding and Metagenomics 10: e172476. <https://doi.org/10.3897/mbmg.10.172476>

Copyright: © Agathe Pirog et al.

This is an open access article distributed under terms of the Creative Commons Attribution

License (Attribution 4.0 International – CC BY 4.0).

Introduction

Once widespread across Europe, the grey wolf *Canis lupus* faced severe persecution during the 19th and 20th centuries, resulting in regional extinctions (Ripple et al. 2014). Since the early 1990s, legal protection and shifting human attitudes have enabled wolf recovery in terms of both range and numbers throughout Europe (Chapron et al. 2014). As wolves recolonise human-dominated landscapes, conflicts with pastoralism intensify, largely due to livestock depredation (Morehouse and Boyce 2017; Bruns et al. 2020; Gervasi et al. 2021; Grente et al. 2022). Therefore, understanding the diet of this species is critical for the development of effective management and conservation strategies.

Accurately determining dietary composition relies on robust methodologies. Traditionally, scat analysis, which involves the mechanical sorting of undigested remains, has been used (Leopold and Krausman 1986; Spaulding et al. 2000). Although informative, this method is prone to well-documented biases (Reynolds and Aebischer 1991; Ciucci et al. 1996; Spaulding et al. 2000; Casper et al. 2007). DNA metabarcoding has emerged as a powerful tool to overcome many of these limitations (Forin-Wiart et al. 2018; de Sousa et al. 2019; Jusino et al. 2019).

In brief, DNA metabarcoding involves: (i) DNA extraction from a sample containing partially or fully digested prey remains; (ii) PCR amplification of short DNA fragments (typically < 300 bp due to DNA degradation) from all organisms present; (iii) high-throughput sequencing (HTS) of the amplified products and (iv) downstream bioinformatics analysis to infer the taxonomic composition of the diet. For this process, the selection of molecular markers is crucial and mitochondrial DNA (mtDNA) is commonly targeted because of its high copy number, which facilitates amplification from degraded samples and its extensive representation in reference databases. Moreover, mitochondrial regions often combine sufficient conservation to allow the design of universal primers with enough sequence variation to discriminate amongst species (Simon et al. 2006; Yang et al. 2014).

Within mtDNA, regions such as cytochrome b and 12S rRNA genes are most commonly employed in carnivore diet studies, as they generally allow species-level identification (Forin-Wiart et al. 2018; Di Bernardi et al. 2021; Massey et al. 2021; Roffler et al. 2021; Shi et al. 2021; Buzan et al. 2024b). However, these markers may be too conserved to differentiate closely-related species, leading to taxonomic assignments that are limited to the genus level (Hacker et al. 2021; Shi et al. 2021; Buzan et al. 2024b). This limitation is particularly evident with *Ovis aries* and *O. musimon*, whose high genetic similarity and complex evolutionary history (Nadler et al. 1973; Rezaei et al. 2010; Meadows et al. 2011; Ciani et al. 2020) render standard markers inadequate for their discrimination (Barbato et al. 2017; Chen et al. 2021).

Yet, distinguishing between *O. aries* and *O. musimon* in carnivore diet analyses is crucial, as these species have markedly different ecological and socio-economic implications. While domestic sheep are closely tied to human-wildlife conflicts, mouflons are a wild species relevant to conservation and game management. Failure to discriminate between them can lead to misinterpretations of predation patterns and flawed management recommendations.

Given these limitations, the mitochondrial D-loop, a highly variable non-coding region, has emerged as a particularly promising candidate for distinguishing the contributions of *O. aries* and *O. musimon* to carnivore diets. Its known utility

in resolving *Ovis* phylogeny and distinguishing subspecies is well established (Hiendleder et al. 1998; Hiendleder et al. 2002; Feng et al. 2009; Buzan et al. 2024a). Moreover, its successful application in carnivore diet studies (Schroeder et al. 2021) and in other closely-related *Ovis* and *Capra* species (Khalatbari et al. 2022) further suggests its potential to overcome the current diagnostic challenge.

To address this critical challenge, the present study aimed to develop and validate a novel, highly discriminatory D-loop marker for the species-level identification of *O. aries* and *O. musimon* in carnivore diet metabarcoding. We achieved this by screening D-loop variation in French reference *O. aries* and *O. musimon* samples and selecting a 178 bp fragment. We subsequently tested the performance of this marker in 115 wolf scat samples collected from the same area, compared its efficacy with a general metabarcoding marker (a portion of the 12S gene) and evaluated different bioinformatic approaches.

Materials and methods

Study area

This study was conducted in France across several Departments in the French Alps. This region represents the core distribution range of the grey wolf in France (Duchamp et al. 2012). In these landscapes, the grey wolf preys primarily on roe deer (*Capreolus capreolus*), red deer (*Cervus elaphus*) and wild boar (*Sus scrofa*) in lowland habitats and on chamois (*Rupicapra rupicapra*) and European mouflon (*O. musimon*) in mountainous areas. This region presents a unique challenge for dietary analysis due to the co-occurrence of wild mouflon and domestic sheep.

European mouflon populations in France originated from introductions, either directly from the Corsican population or through secondary dispersal. The Alps form the core of the national distribution, accounting for 63% of its range. Smaller populations are also present in the Massif Central (18%) and Pyrenees (17%). As of 2016, mouflons occupied approximately 5,084 km² across the country (Savouré-Soubelet et al. 2021).

Additionally, approximately 1.5 million sheep are bred annually in the French Alps, mainly for meat and dairy production (IDELE 2018). During summer, a substantial proportion of these herds engage in transhumance, migrating to high-altitude Alpine pastures. Wolf depredation on livestock is reported each year, with recurring hotspots, where conflicts could lead to economic losses (Linnell and Cretois 2018; Gervasi et al. 2021; Grente et al. 2022).

The presence of both wild mouflon and domestic sheep creates a complex dietary landscape for wolves, directly leading to the challenge of differentiating these two species in diet analyses.

Sampling

Ovis reference samples

A comprehensive set of reference samples is required to develop a robust discriminatory marker.

Initial searches of the GenBank public database (<https://www.ncbi.nlm.nih.gov/>) for *Ovis aries* and *Ovis musimon* D-loop sequences revealed limitations:

none of the available *O. musimon* D-loop sequences ($n = 27$) originated from French populations, while none of the numerous *O. aries* D-loop sequences with associated geographic information corresponded to breeds commonly raised in France. Moreover, no haplotype identified in the database was exclusive to either species, highlighting the lack of species-specific resolution in the available data.

To address this gap, we collected samples of both species within the study area. Specifically, for *O. aries*, tissue samples were obtained from the main meat processing plant in the Alps, which processes 88% of sheep production in the Alpes-Provence Region. The samples included various breeds, such as Merinos, Lacaune, Merinos $\times$ Île-de-France, Mourerous, Préalpes and Préalpes $\times$ Merinos crosses (Table 1). For *O. musimon*, samples were either collected from legally hunted or captured animals (tissue, $n = 12$; blood, $n = 4$) or non-invasively (faeces, $n = 20$; hairs, $n = 10$) from the Corsican population and the four main populations of south-western France: Mercantour and Cadarache (Alps), Caroux and Aigoual (Massif Central) (Table 1).

Table 1. Reference *Ovis musimon* and *Ovis aries* samples collected in France. N_{sample} : number of samples collected within each population or breed. IDF: Ile-de-France.

<i>Ovis musimon</i>		<i>Ovis aries</i>	
Population	N_{sample}	Breed	N_{sample}
Corse	10	Lacaune	1
Cadarache (Alps)	4	Merinos	11
Mercantour (Alps)	12	Merinos X IDF	1
Aigoual (Massif Central)	8	Mourerous	6
Caroux (Massif Central)	12	Préalpes X Merinos	1
		Préalpes	12
		Unknown	14

In total, 46 sheep and 46 European mouflon samples were collected. These samples were subsequently used to sequence the entire D-loop region to identify a short diagnostic fragment for species-level discrimination in metabarcoding studies.

Wolf scat samples

Since the return of the grey wolf to France in the early 1990s, the population has been closely monitored by French authorities. Non-invasive samples, primarily scats, have been collected for over 30 years and analysed using various molecular markers to confirm species identity and perform individual identification following a standardised high-quality workflow (Pirog et al. 2025).

Amongst these, a subset of samples collected in the same geographic area as the reference *Ovis* samples and previously analysed by metabarcoding revealed the presence of *Ovis* spp. prey in 115 cases (unpublished data). These 115 samples (Fig. 1) were selected for the present study to: (i) confirm the presence of *Ovis* spp. through metabarcoding sequencing of a portion of the 12S gene and (ii) evaluate the discriminatory power of the diagnostic D-loop fragment identified from *Ovis* reference samples.

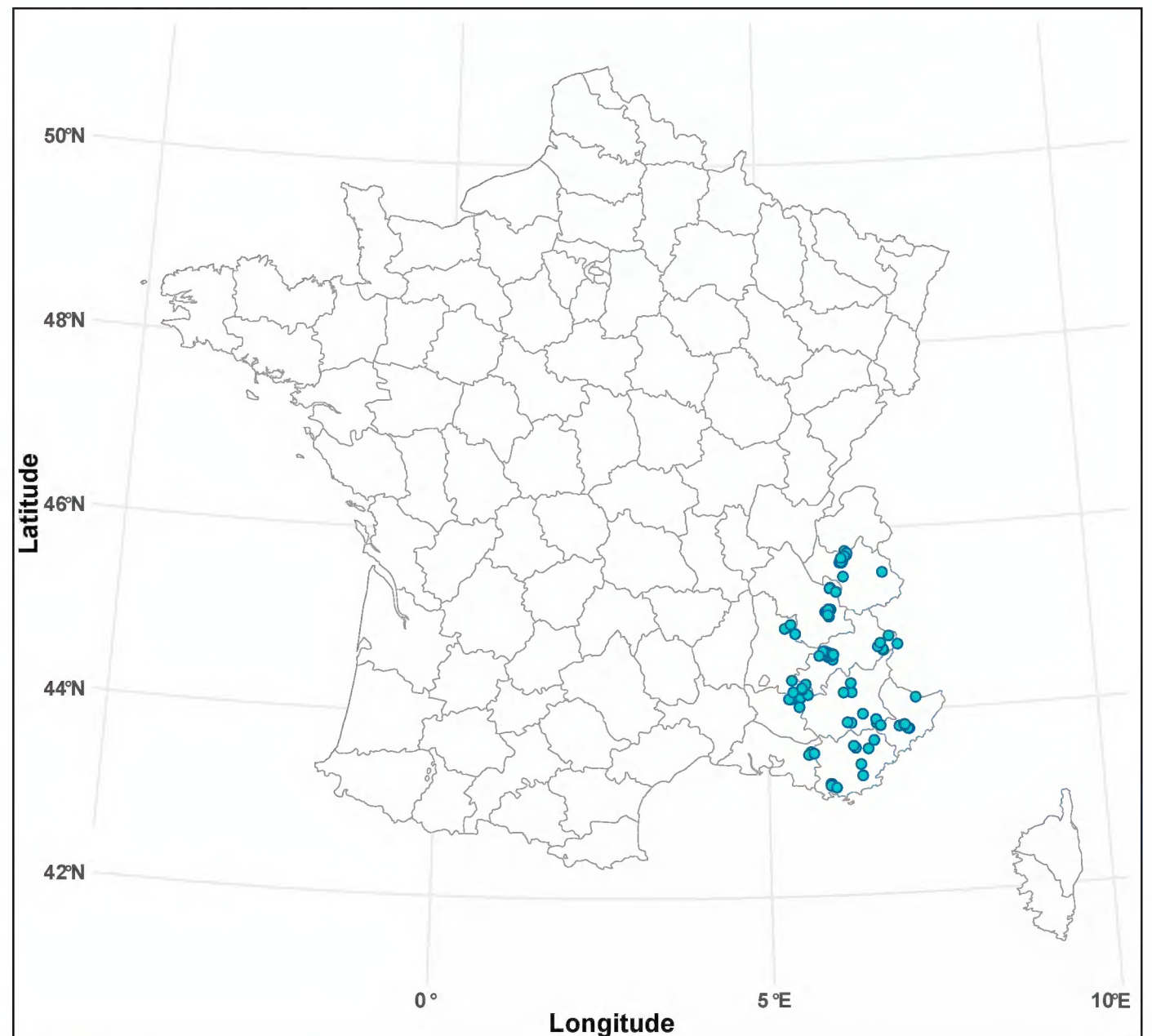


Figure 1. Geographic distribution of 115 wolf scat samples used for metabarcoding.

Laboratory procedures

General quality control procedures

For all laboratory procedures, that is, sequencing of *Ovis* reference samples and metabarcoding of wolf scat samples, a unidirectional workflow was followed, with dedicated spaces for rare and degraded DNA at each step. DNA extraction was performed separately for tissue or blood and non-invasive samples using an ultraviolet light purification system in a room dedicated to non-invasive samples. The extracted DNA samples were stored at -20 °C. Sensitive reagents (enzymes and DNA primers) used for DNA amplification were prepared in a clean room with positive air pressure. DNA and reagents were then assembled (manual and automated work) in a separate room, while amplification and sequencing were performed in another room with a negative air pressure. Sterile disposable tools were used to manipulate samples and all workstations were cleaned with bleach after each step (and each sample during extraction) to avoid cross-contamination.

Ovis reference samples

Each tissue, blood or hair bulb sample was transferred to a sterile and labelled microtube for DNA extraction. For faeces preserved in ethanol, the supernatant was first transferred to a numbered microtube, centrifuged to remove ethanol and DNA was extracted from the pellet. Two negative extraction controls

(blanks) and one positive extraction control (sample previously analysed and validated for DNA quality) were added to a 96-well plate. Samples were lysed with positive and negative extraction controls according to the manufacturer's instructions (NucleoSpin 96 Tissue Kit and NucleoSpin 96 Blood Kit; Macherey-Nagel, Düren, Germany). DNA was isolated and purified using purification columns and vacuum filtration methods.

The mtDNA D-loop was then amplified and sequenced using in-house designed primers *Ovis*-CR-longread-F (5'-AAACTCCCAAACATACAACACGG-3') and *Ovis*-CR-longread-R (5'-TGTATRTGACCCAGGTGCCT-3'), which amplify a 1,100 bp long fragment.

PCR amplification was performed in a 96-well microplate with a final volume of 10 µl containing 5 µl LongAmp Hot Start Master Mix Taq (New England Biolabs, Ipswich, MA, USA), 0.40 µM non-fluorescent CR primer pairs and 2 µl DNA extract. The thermal cycle was as follows: initial denaturation at 94 °C for 30 s, followed by 31 cycles of 94 °C for 30 s, 61 °C for 40 s and 65 °C for 1 min, followed by a final extension at 65 °C for 10 min. Three control PCR blanks and three positive control DNA samples were added to the microplate.

The PCR products were bidirectionally sequenced using the Sanger method with the same primers and a BigDye Terminator v.3.1 Cycle Sequencing Kit (Applied Biosystems, Foster City, CA, USA).

After purification, the sequences were analysed using an ABI PRISM 3130 XL capillary sequencer (Applied Biosystems, Foster City, CA, USA) and the electropherograms were interpreted using Geneious software (Kearse et al. 2012).

Wolf scat samples

Although the general methodology of metabarcoding is well established, specific choices in laboratory protocols and bioinformatic pipelines vary widely (Alberdi et al. 2018; Ando et al. 2020; Bohmann et al. 2022; Hakimzadeh et al. 2024), necessitating careful consideration. Therefore, we explain our chosen methods following best practice guidelines (Klymus et al. 2024) to ensure clarity and reproducibility. The DNA metabarcoding workflow is illustrated in Fig. 2.

To optimise resource allocation and given our primary objective to evaluate the ability of the marker for *Ovis* species-level assignments rather than individual diet assessment, we prioritised biological replicates over technical replicates for each sample. We thus performed one replicate per sample to allow for metabarcoding of a larger number of samples (Ando et al. 2020).

First, DNA was extracted from wolf scat samples following the protocol described by Pirog et al. (2025). Predator species and individual identification were performed by sequencing the mitochondrial D-loop region and genotyping 22 microsatellite markers, as detailed in the same reference.

Metabarcoding analysis was subsequently performed by amplifying two fragments: a portion of the 12S rRNA gene sequence to confirm the presence of *Ovis* spp. prey using this marker and the diagnostic D-loop fragment specific to the *Ovis* species. Each targeted region was amplified in different 96-well plates to avoid inversion or barcoding errors. Negative PCR controls (n = 13) as well as eight positive PCR controls consisting of brown bear *Ursus arctos* hair samples were added to the samples.

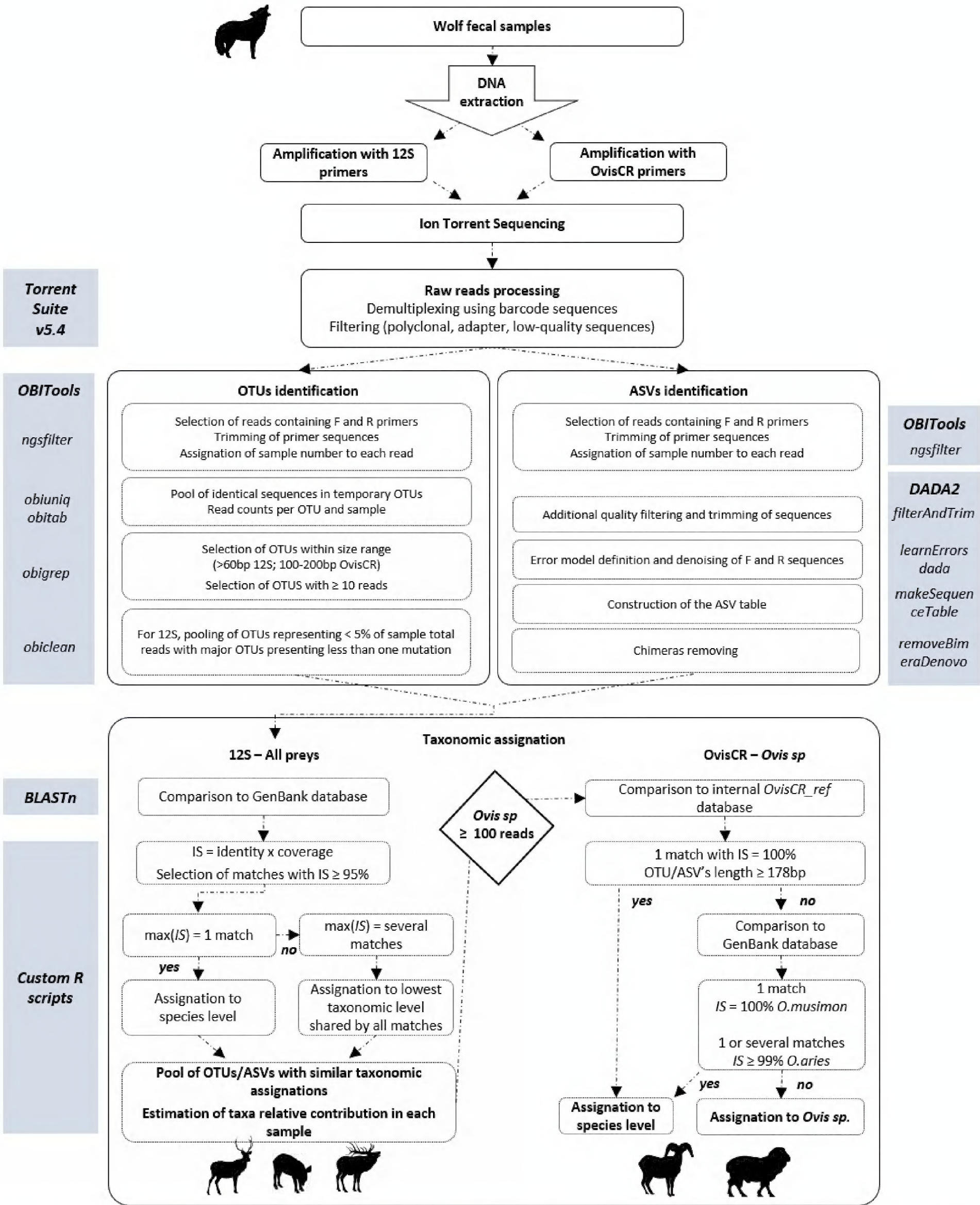


Figure 2. Metabarcoding pipeline used to identify prey in grey wolf faecal samples. The grey boxes indicate the software used and the specific functions are indicated in *italics*. OTU: Operational taxonomic unit; ASV: Amplicon sequence variant; F: Forward; R: Reverse; IS: Identity score.

For the 12S region, primers Vert-NGS-12S-F (5'-GGATYAGATACCCYRY-TATG-3') and Vert-NGS-12S-R (5'-TRGRACAGGCTCCTCTRR-3'), adapted from 12S_V5 primers (Riaz et al. 2011), were used. Furthermore, an

in-house designed blocker to avoid amplification of grey wolf DNA was used (5'-CTATGCTTAGCCCTAAACATAGATAATTTTACAACAAAATAATTC-3').

Primers for the diagnostic D-loop fragment (forward: 5'-CCYCAYGCAATATAAGCA-3'; reverse: 5'-GTGAACAAGCTCGTGATC-3') were newly designed. All primers were modified to include Ion Torrent-compatible adapters and sample-specific barcodes according to the Ion Amplicon Library Preparation Fusion Method (Thermo Fisher Scientific, Waltham, MA, USA), in a one-step PCR. While increasing costs, a one-step PCR has been shown to reduce risk of cross-contamination and tag-jumps (Bohmann et al. 2022). Sample-specific barcodes were only added in front of the forward primers as sequencing was only performed in one direction. While potentially reducing the quality of sequences at the end, amplicons targeted were of size well below the threshold-length of 400 bp allowed by the library sequencing protocol and sequencing in one-direction should thus not impact end-reads quality.

Libraries were prepared following the Ion Amplicon Library Preparation Fusion Method protocol, provided by Thermo Fisher Scientific, with some adaptations.

PCR amplification was performed in 10 µl reactions containing 5 µl of NEB Next Ultra II Q5 Mastermix (New England Biolabs, Ipswich, MA, USA), 0.25 µM of each primer, 2.50 µM of the grey wolf blocker for the 12S fragment and 3 µl template DNA. The cycling conditions were as follows: initial denaturation at 98 °C for 3 min; 40 cycles of 98 °C for 10 s, 53 °C for 30 s and 72 °C for 30 s; and final extension at 72 °C for 5 min.

PCR products were quantified by qPCR (Suppl. material 1), equimolarly pooled between samples (both targets included) and then purified with the Agencourt AMPure XP bead system (Beckman Coulter) with a 1.5× ratio of beads.

All libraries were qualified using a QIAxcel Advanced instrument following the manufacturer's protocol for the High Resolution Kit and their concentrations were assessed using digital PCR (Suppl. material 1). Libraries were diluted and pooled to achieve a 35 pM equimolar concentration for the template reactions. This process involved attaching DNA fragments to Ion Sphere Particles (ISPs) for clonal amplification via emulsion PCR, conducted using the Ion Chef™ System (Thermo Fisher, Waltham, MA, USA). Following preparation, the samples were sequenced on the Ion S5™ XL system using an Ion Chef 530 Kit Chef (Thermo Fisher Scientific, Waltham, MA, USA).

Analyses

Ovis reference samples

D-loop sequences obtained from the *Ovis* reference samples were processed using Geneious (Kearse et al. 2012). Forward and reverse reads were aligned for each sample using Geneious alignment to generate high-quality consensus sequences.

These consensus sequences were aligned in Seaview v.4 (Gouy et al. 2009) and trimmed at both the 5' and 3' ends to obtain a standardised 674 bp fragment. Distinct haplotypes were identified using DNA Collapser (<https://users-birc.au.dk/palle/php/fabox/dnacollapser.php>) and manually inspected to locate a 100–300 bp region that: (i) discriminated between *O. aries* and *O. musimon* and (ii) contained sufficiently conserved flanking regions for primer design. Haplotype diversity indices were calculated in R (R Core Team 2023)

using the package *pegas* (Paradis 2010). Haplotype rarefaction curves were generated for each species using 1,000 iterations.

A TCS haplotype network (Clement et al. 2000) was constructed using PopArt v.1.7 (Leigh et al. 2015) to visualise genetic relationships amongst samples. All identified haplotypes were compiled into a reference database (*OvisCR_ref*) for downstream identification and comparison with publicly available sequences.

Wolf scat samples

To assess the performance of the newly-developed D-loop marker and compare it with the 12S marker in actual wolf scat samples, we processed metabarcoding data using two distinct bioinformatic pipelines.

Indeed, multiple pipelines have been described to analyse metabarcoding data, with key steps including demultiplexing raw reads, quality filtering, merging paired-end reads, artefact filtering, clustering sequences in operational taxonomic units (OTUs) or defining amplicon sequence variants (ASVs) and, finally, taxonomic assignments (Hakimzadeh et al. 2024). Some studies advocate the use of thresholds to filter reads in low numbers and cluster remaining sequences in OTUs, but setting these thresholds remains arbitrary with no consensus method amongst studies (Drake et al. 2022; Littleford-Colquhoun et al. 2022; Tercel and Cuff 2022). By denoising sequences before constructing OTUs, ASV methods do not rely on setting dissimilarity thresholds and may be more accurate in identifying taxa (Callahan et al. 2017; Couton et al. 2021).

Here, we tested two commonly used pipelines to define: (i) OTUs using common dissimilarity thresholds with the OBITools package v. 1.01 (<https://python-hosted.org/OBITools/>; Boyer et al. (2016)) and (ii) ASVs using DADA2 R package v. 1.26.0 (Callahan et al. 2016) to assess their impacts on the results. The parameters and main functions are detailed in Suppl. material 2. Raw read FASTQ files are available in the European Nucleotide Archive (ENA) database (Project Accession PRJEB106338) to allow readers to test other pipelines. All analyses were performed with R v. 4.2.2. OTUs and ASVs were identified following the same approach for both the 12S and diagnostic D-loop (*OvisCR*) fragments.

First, raw reads generated by Ion Torrent sequencing were processed using Torrent Suite™ v.5.4 (Thermo Fisher Scientific, Waltham, MA, USA). Reads were demultiplexed according to the barcode sequences and polyclonal, low-quality and adapter-contaminated sequences were filtered out using the default quality parameters. One FASTQ file per sample was generated, containing all the retained sequences for both targeted regions.

To define OTUs, the resulting reads were analysed using the OBITools package. Only sequences containing both primers (forward and reverse) for each targeted region were retained and the primers were trimmed from the sequences. Sample identification numbers were assigned to each read. Strictly identical sequences were grouped into temporary OTUs and merged to compute the read counts per OTU. Subsequently, only OTUs of the expected size range (> 60 bp for 12S and 100–200 bp for *OvisCR*) and with a minimum of 10 reads were retained for further analysis. This threshold was chosen, based on OBITools recommended parameters (Royaux et al. 2022) and previous metabarcoding studies (Drake et al. 2022), in order to reduce spurious low-frequency variants while retaining rare, but genuine prey sequences

(Deagle et al. 2013). For the 12S target, OTUs representing less than 5% of the total reads in a sample and differing by only a single mutation (97% similarity threshold, Alberdi et al. (2018)) from a major OTU were merged with the corresponding major OTU, as they were considered likely artefacts. This step was not applied to the OvisCR target because of the close genetic similarity between *O. musimon* and *O. aries* haplotypes. Such merging could inadvertently obscure the diagnostic differences that this study aims to resolve. The final datasets consisted of one table per target, listing all validated OTUs and their relative abundances per sample.

In parallel, ASVs were defined using the DADA2 R package. First, OBITools was used to filter sequences containing both primers (forward and reverse) for each targeted region and to trim the primers from the sequences, as this step was not performed by DADA2. Only reads with no uncertain base (no Ns) and containing less than three expected errors were retained. We then used the DADA2 algorithm to learn the error model from the data and applied it to the filtered and trimmed sequence data. No additional read-count threshold was applied for this pipeline, as the denoising algorithm explicitly models sequencing errors. Finally, an ASV table was constructed for each target before removing the chimeras. The final datasets consisted of one table per target, listing all validated ASVs and their relative abundances in each sample.

Finally, the identified OTUs or ASVs were compared to internal reference databases or the GenBank public database for taxonomic assignment using BLASTn and custom R scripts. The approach varied depending on the target fragment.

For the 12S fragment, OTUs/ASVs were aligned with the GenBank public database. For each alignment, an identity score (*IS*), defined as the product of the sequence identity and alignment coverage, was calculated. Only matches with an $IS \geq 95\%$ were considered. Taxonomic assignment followed a conservative approach: when a single best match was identified, the OTU/ASV was assigned to that species. In cases where multiple top matches were identified with the same *IS*, the OTU/ASV was assigned to the lowest shared taxonomic level amongst the matches. Finally, reads from OTUs/ASVs with identical taxonomic assignments were pooled to estimate the relative contribution of each taxon in each sample to diet composition.

For samples with more than 100 *Ovis* spp. reads, based on the 12S fragment, taxonomic assignments were refined using the OvisCR dataset. In the first step, the OvisCR OTUs/ASVs were aligned against the internal reference database (*OvisCR_ref*). When a perfect match of at least 178 bp (the length of OvisCR) was found ($IS = 100\%$), the haplotype was confirmed and the species was assigned accordingly. In the second step, OTUs/ASVs that did not meet these criteria were compared with GenBank sequences. At this stage, species-level assignment was accepted if the OTU/ASV matched an *O. musimon* sequence with $IS = 100\%$ or an *O. aries* sequence with $IS \geq 99\%$. This threshold accounts for the higher mitochondrial diversity in domestic sheep, a result of artificial selection, which increases the likelihood of encountering unreferenced, but authentic haplotypes. In contrast, the lower genetic diversity of *O. musimon* makes undocumented haplotypes less probable. Therefore, OTUs/ASVs matching *O. musimon* with $IS < 100\%$ or matching multiple *Ovis* species with $IS > 99\%$ were conservatively assigned at the genus level.

Results

Ovis reference samples

From the 46 domestic sheep and 46 European mouflon samples, high-quality D-loop sequences (674 bp) were obtained for 41 (89% of the *O. aries* samples collected) and 30 samples (65% of the *O. musimon* samples collected), respectively. Most of the *O. musimon* samples with low quality sequences ($n = 16$) were non-invasive samples, comprising 14 scat or hair samples, with degraded or low quantity DNA. This may explain the lower sequencing success in *O. musimon* compared to *O. aries*. A total of 45 distinct haplotypes (GenBank accession numbers [PX219566–PX219610](#)) were identified with 82 variable sites and no overlap between species. Haplotype diversity (Hd) was notably higher in domestic sheep, with 40 haplotypes identified and a haplotype diversity of 0.999, than in European mouflon, with five haplotypes identified and a haplotype diversity of 0.685.

From this alignment, a 178 bp fragment suitable for metabarcoding was identified (OvisCR), with no haplotypes shared between the two species and flanking regions sufficiently conserved to design primers with minimal use of degenerate bases. It included 43 variable sites and discriminated between four European mouflon ($Hd = 0.669$) and 35 domestic sheep haplotypes ($Hd = 0.985$; Fig. 3 and Suppl. material 3). Most European mouflon haplotypes differed from domestic sheep haplotypes by at least three mutation sites. An exception was CR178_OM004 (shared by two individuals and located at the centre of the haplotype network in Fig. 3), which differed by only one mutation site from two domestic sheep haplotypes. On the longer CR fragment (674 bp), this haplotype differed by at least six mutations from the closest domestic sheep haplotype and was not at the centre of the network (Suppl. material 4).

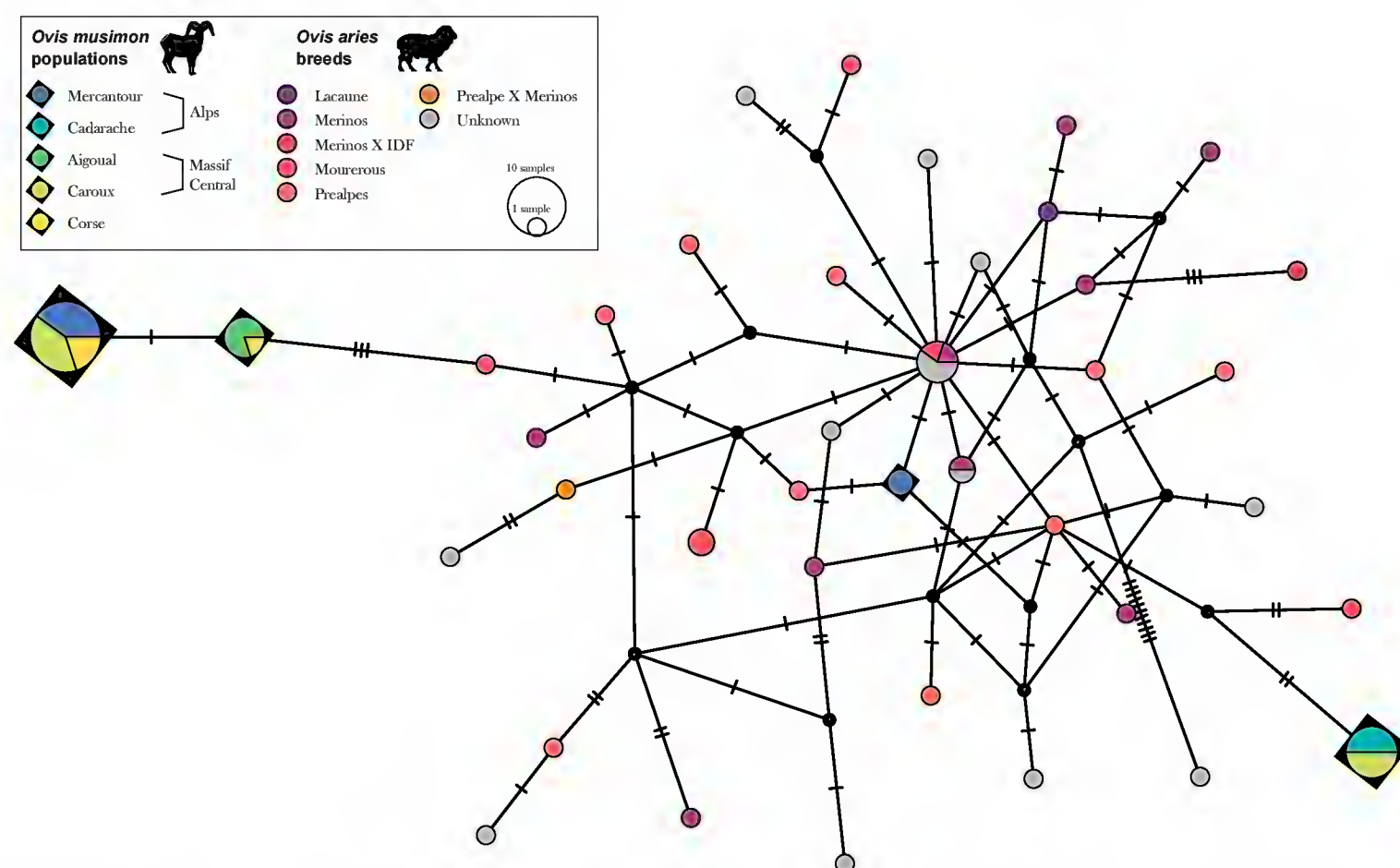


Figure 3. TCS haplotype network, based on a 178 base pair diagnostic D-loop fragment, constructed from 30 *Ovis musimon* and 41 *Ovis aries* reference samples collected in France. Each haplotype is represented by a circle for *O. aries* and a circle enclosed within a square for *O. musimon*. Each tick mark on the lines connecting the haplotypes indicates a single mutation. The circle size is proportional to the number of individuals sharing that haplotype and the colours indicate the breeds of *O. aries* and populations of *O. musimon*. Black dots represent the unsampled haplotypes. IDF, Ile-de-France.

The higher genetic diversity in domestic sheep was also highlighted in the haplotype rarefaction curves, with the number of detected haplotypes increasing almost linearly with sample size (Fig. 4). In contrast, the curve for the European mouflon reached an early asymptote, indicating that further sampling is unlikely to reveal additional haplotypes.

Due to the high haplotypic diversity observed in *O. aries* and the close genetic proximity between haplotypes of the two species, as revealed by the TCS network using the OvisCR fragment, we conducted BLASTn searches against the GenBank database to assess the taxonomic specificity of each haplotype. Only matches with 100% query coverage and 100% identity were considered.

Amongst the 35 haplotypes identified in *O. aries* samples, 19 had exact matches in GenBank, all corresponding exclusively to *O. aries* sequences.

All four haplotypes detected in the *O. musimon* samples also yielded matches. One haplotype (CR178_OM002) matched sequences from *O. musimon*, *O. ammon* and *O. orientalis*. However, as the latter two species are not present in France, this does not compromise the ability of the OvisCR fragment to identify *O. musimon* prey in our study area. Another haplotype (CR178_OM004) matched one *O. ammon* sequence and 16 *O. aries* sequences, all originating from local or rare breeds not present in France (e.g. from Spain, Portugal, Italy, Bulgaria, Hungary, Poland, Russia, Turkiye, Pakistan, China and Mongolia). Nevertheless, because of its close proximity to other French *O. aries* haplotypes and its central position in the network, we conservatively classified the CR178_OM004 haplotype detected in wolf scat as *Ovis* spp. prey, without specifying the species.

The remaining *O. musimon* haplotypes matched exclusively with *O. musimon* sequences.

Wolf scat samples

Negative and positive controls were used to determine whether additional filtering using read count thresholds should be used when interpreting metabarcoding results.

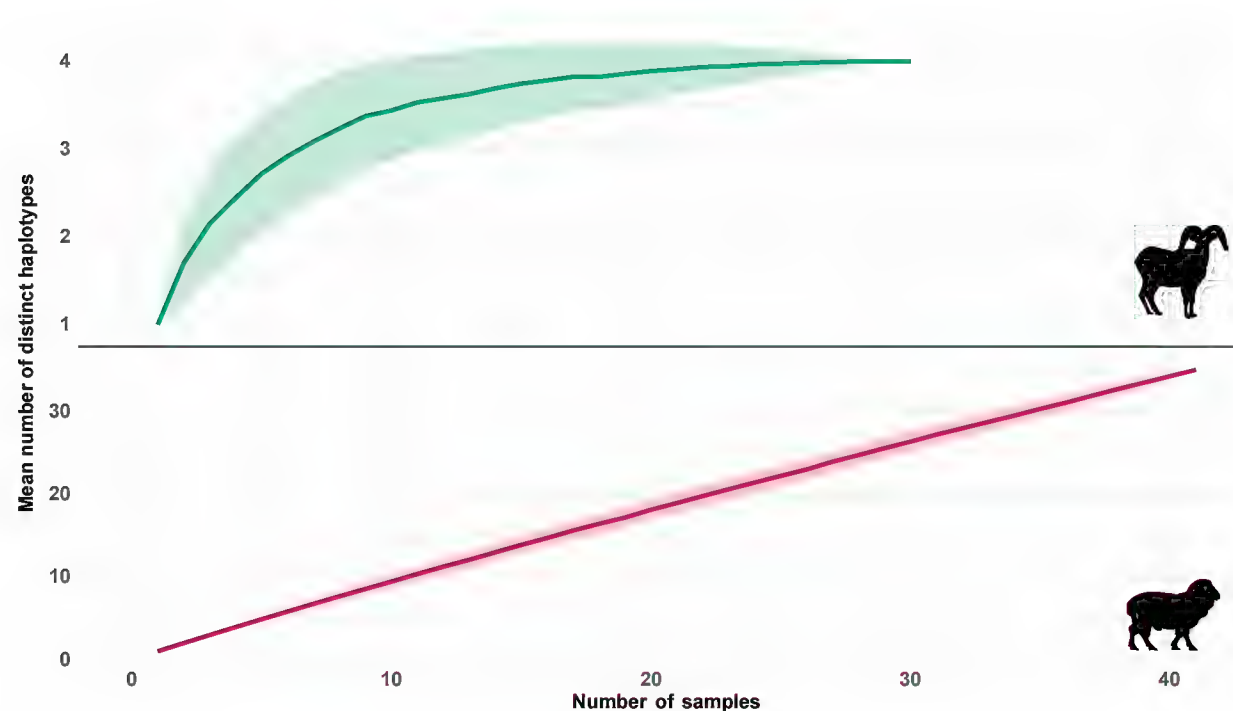


Figure 4. Haplotype rarefaction curves, based on a 178 base pair diagnostic D-loop fragment generated from 30 *Ovis musimon* and 41 *Ovis aries* reference samples collected in France. The curves show the accumulation of unique haplotypes with increasing sample sizes for each species. The shaded areas represent the standard deviations based on 1,000 re-sampling iterations. Green, *Ovis musimon*; Purple, *Ovis aries*.

Regarding the 12S fragment, only two negative controls exhibited no reads using either the OBITools or DADA2 pipeline (Suppl. material 5). Other negative controls exhibited one to 345 reads and the two controls exhibiting the highest number of reads (82 and 219 with OBITools and 95 and 345 with DADA2) were contaminated with human DNA. Positive controls exhibited high numbers of *U. arctos* reads (between 14,861 and 110,003 for OBITools and between 16,666 and 127,212 for DADA2), as expected. Other species were sometimes detected with a number of reads varying from one to 77 with OBITools and from three to 166 with DADA2.

Regarding the OvisCR fragment, one negative control exhibited one *Ovis* sp. read, with either OBITools or DADA2. One positive control also exhibited five *O. aries* reads using the CR fragment and DADA2 pipeline.

These results highlight the need to apply read-count thresholds when interpreting metabarcoding results, especially for the 12S fragment. Based on both control samples and thresholds commonly used in literature, we applied a threshold of 100 reads in a sample to consider an OTU/ASV as a true result and not an artefact or a spurious sequence. This threshold was applied using either OBITools or DADA2.

For the OvisCR fragment, a less restrictive threshold may be used, as controls rarely exhibited *Ovis* spp. reads and when it was the case, it was never above 10 reads. This value is consistent with thresholds previously applied in other metabarcoding studies. Thus, we discuss the results for this fragment using both 10 reads and 100 reads thresholds per OTU/ASV per sample.

Using the 12S fragment, the presence of *Ovis* spp. prey was confirmed in all 115 wolf samples, with read counts ranging from 108 to 45,973 (mean: 4,294 reads) with OBITools and from 144 to 75,743 (mean: 10,380 reads) with DADA2 (Suppl. material 5).

With the OvisCR fragment, *Ovis* spp. prey was detected with more than 10 reads in all but five samples (95.7%) with OBITools and in all but four samples (96.5%) with DADA2, with read counts per sample ranging from one to 54,161 (mean: 4,303 reads) with OBITools and from one to 6,478 (mean: 4,623 reads) with DADA2 (Suppl. material 5). All samples, except 13 (88.7%) and 10 (91.3%), had more than 100 *Ovis* spp. reads with OBITools and DADA2, respectively.

Samples with fewer than 100 *Ovis* spp. reads with the OvisCR fragment also exhibited significantly fewer *Ovis* spp. reads in the 12S dataset than the other samples (OBITools: mean = 1,275 vs. 4,679; t -test, $p = 4.06 \times 10^{-4}$; DADA2: mean = 2,213 vs. 11,157; t -test, $p = 2.84 \times 10^{-8}$), which may explain the low detection with the OvisCR fragment.

When considering only samples with more than 100 *Ovis* spp. reads with OvisCR, 85 and 91 were assigned at the species-level with OBITools and DADA2, respectively (Table 2). Samples in which *Ovis* spp. prey identification was only possible at the genus level had significantly fewer reads than those with species-level assignments (OBITools: mean = 1,917 vs. 5,353 reads; t -test, $p = 1.33 \times 10^{-3}$; DADA2: mean = 2,394 vs. 5,394 reads; t -test, $p = 8.35 \times 10^{-3}$). Finally, at this threshold, the two pipelines yielded different assignments for eight samples, of which seven were assigned at the species-level with DADA2, but not with OBITools and one was assigned at the species-level with OBITools, but not with DADA2.

When considering samples with more than 10 *Ovis* spp. reads with OvisCR, 96 and 100 were assigned at the species-level with OBITools and DADA2,

Table 2. Number and percentage of wolf scat samples (n = 115) with *Ovis* prey successfully assigned at the species-level using the OvisCR fragment. Results are shown for two bioinformatic pipelines (OBITools and DADA2) and two read counts thresholds (10 and 100 reads). Percentages are presented relative to the total number of samples and relative to samples with *Ovis* reads above the threshold.

Read count threshold	Pipeline	Number of samples assigned at species-level	% of total samples	% of samples with <i>Ovis</i> reads above threshold
100 reads	OBITools	83	72%	81%
	DADA2	91	79%	87%
10 reads	OBITools	96	83%	87%
	DADA2	100	87%	90%

respectively (Table 2). Samples in which *Ovis* prey identification was only possible at the genus level had significantly fewer reads than those with species-level assignments (OBITools: mean = 1,508 vs. 5,238 reads; *t*-test, $p = 1.56 \times 10^{-4}$; DADA2: mean = 1,603 vs. 5,366 reads; *t*-test, $p = 2.18 \times 10^{-4}$). Finally, at this threshold, the two pipelines yielded different assignments for 10 samples, of which seven were assigned at the species-level with DADA2, but not with OBITools and three were assigned at the species-level with OBITools, but not with DADA2. Amongst the samples with different assignments, two contained ASVs from both species with DADA2, one assigned as *O. musimon* by OBITools (containing 1508 *O. musimon* reads and 11 *O. aries* reads with DADA2) and the other assigned as *Ovis* sp. by OBITools.

Overall, these results indicate that species-level assignment reliability increases with read depth and that lower thresholds mainly affect samples with limited prey DNA, rather than introducing systematic mis-assignments. Furthermore, DADA2 consistently assigned a slightly higher number of samples at the species-level than OBITools, particularly at lower read thresholds, although both pipelines converged towards similar conclusions when read depth was sufficient.

Discussion

We developed and validated a novel D-loop mitochondrial fragment to differentiate between domestic sheep (*O. aries*) and European mouflon (*O. musimon*) in diet metabarcoding studies. This marker directly addresses a critical limitation of commonly used metabarcoding markers (e.g. 12S or cytochrome b) which often lack the resolution to distinguish closely-related species, particularly within the genus *Ovis*, in carnivore diets. Accurate discrimination between these species is particularly important in the context of large carnivore ecology, where misidentification can lead to biased estimates of livestock depredation and misinformed management decisions.

Identification of a diagnostic fragment for *Ovis* species discrimination

By sequencing the entire D-loop region in 92 reference samples collected in France, we identified a 178 bp fragment, herein named OvisCR, capable of discriminating between the two species despite their close evolutionary relationship. Most haplotypes were separated by at least three mutation sites, allowing reliable species-level discrimination in the majority of cases. Rare instances of minimal divergence, such as the CR178_OM004 haplotype

identified in *O. musimon*, but only one mutation away from several *O. aries* haplotypes, underscore both the evolutionary proximity of the two species (Barbato et al. 2017; Chen et al. 2021) and the inherent limitations of short metabarcoding fragment.

The OvisCR fragment also revealed significantly higher genetic diversity in domestic sheep (35 haplotypes) compared to European mouflon (4 haplotypes), consistent with previous studies (Lawson Handley et al. 2007; Rochus et al. 2018). This pattern likely reflects the artificial selection and diverse geographic origins of *O. aries* breeds (Ryder 1991), in contrast to founder effects and restricted gene flow in introduced *O. musimon* populations (Savouré-Soubelet et al. 2021 Portanier et al. 2022). Rarefaction curves confirmed that all European mouflon CR haplotypes within our study area were identified with a relatively small sample size, whereas additional sampling of domestic sheep would likely reveal further haplotypic diversity.

Finally, some *O. musimon* CR haplotypes matched sequences from other *Ovis* species and from specific regional *O. aries* breeds, which must be considered in areas where these species are also present. Furthermore, 40% of the *O. aries* haplotypes detected in our study were not referenced in public databases which would have resulted in genus-level assignment had we relied solely on public resources. It may be due to the genetic distinctiveness of French breeds compared to other European or Asian breeds referenced in public databases (Rochus et al. 2018). Therefore, we strongly recommend that metabarcoding studies consider local livestock diversity and integrate regional reference datasets when necessary before interpreting matches.

Application of the OvisCR fragment to wolf scat samples

Application of the OvisCR fragment to 115 wolf scat samples demonstrated its practical value in an empirical metabarcoding context. Species-level identification was achieved for more than 81% of the samples with *Ovis* reads, regardless of the pipeline or the threshold used. Samples assigned only at the *Ovis* genus-level were characterised by significantly lower read counts, suggesting that taxonomic resolution was partly constrained by sequencing depth. To ensure the robustness of this application, we further evaluated the performance of the OvisCR fragment across different read count thresholds and bioinformatic pipelines.

The use of read-count thresholds is a critical, yet often arbitrary step in metabarcoding analyses (Drake et al. 2022; Littleford-Colquhoun et al. 2022; Tercel and Cuff 2022). In this study, we removed all OTUs represented by fewer than 10 reads across all samples when using the OBITools pipeline, as this approach is known to retain spurious sequences generated during PCR or sequencing. Such an initial filtering step was not required for DADA2, because its denoising algorithm explicitly models sequencing errors and effectively removes low-abundance artefactual sequences. Following taxonomic assignments, we used control samples to assess whether additional filtering was necessary. While a conservative threshold of 100 reads per OTUs/ASV per sample was justified for the universal 12S marker, based on control data, contamination levels for the OvisCR fragment were consistently low, supporting the exploration of a lower threshold of 10 reads. Lowering this threshold increased the number

of samples assigned at the species-level without introducing clear mis-assignments. Importantly, the impact of read-count thresholds depended more strongly on the molecular marker itself than on the choice of bioinformatic pipeline. Both OBITools and DADA2 showed similar sensitivity to threshold variation for a given marker, whereas threshold effects differed markedly between the universal 12S marker and the taxon-specific OvisCR fragment. This pattern contrasts with the metabarcoding literature (Callahan et al. 2017; Couton et al. 2021) and warrants further investigation using additional datasets.

Comparison of two widely-used bioinformatic pipelines, OBITools and DADA2, showed consistent results when sufficient read depth was available. DADA2 consistently yielded slightly more species-level assignments, particularly at lower read thresholds, likely due to its denoising strategy and explicit error modelling (Callahan et al. 2017; Couton et al. 2021). In contrast, the OTU-based approach implemented in OBITools appeared more conservative for samples with limited read depth. Importantly, discrepancies between pipelines were rare and primarily concerned cases with limited OvisCR reads, rather than conflicting species assignments. It highlights the robustness of the OvisCR marker across different analytical frameworks.

Furthermore, the classification of the CR178_OM004 haplotype illustrates the importance of conservative decision-making in metabarcoding-based inference. Although full-length D-loop data support its affiliation with *O. musimon*, ambiguity at the short-fragment scale led us to deliberately assign this haplotype to *Ovis* sp. rather than to a species. This choice prioritises robustness and transparency over maximal resolution and avoids overestimating wild prey consumption in a management-sensitive context. Such conservative practices are particularly important when closely-related wild and domestic taxa co-exist.

Finally, although the results presented here demonstrate the effectiveness of the OvisCR fragment in a specific ecological and geographic context, caution is warranted when extrapolating these findings to other species or regions. This study does not encompass the full mitochondrial control region diversity of either domestic sheep or European mouflon across other regions of Europe and Asia and the diagnostic value of the OvisCR fragment should, therefore, not be assumed to be directly transferable beyond the French context.

Conclusions

From a conservation and management perspective, the ability to discriminate between predation on *O. aries* and *O. musimon* is essential. Domestic sheep predation is a central driver of human-wolf conflict in France and elsewhere (Linnell and Cretois 2018; Gervasi et al. 2021; Grente et al. 2022), while predation on European mouflon, though relevant for game management, has different implications. Misclassifying one species as the other could lead to incorrect estimates of livestock depredation, misinformed conflict mitigation strategies and flawed compensation or culling decisions.

Our approach can serve as a model for similar developments targeting other taxonomically challenging prey groups. Future work should aim to extend the method to a broader set of herbivore taxa and validate its applicability across diverse geographic regions.

Acknowledgements

The authors thank all volunteers and coordinators of the French Wolf-Lynx network for data collection and fieldwork. We also thank the ANTAGENE technicians who carried out laboratory work and electropherogram readings. We thank Sophie Verzelloni for the administrative support. We finally thank Florence Quirié-Guimbard, Dominique Gauthier and Fanny Bastien from the veterinary services for their valuable contributions to sheep data acquisition.

Additional information

Conflict of interest

The authors have declared that no competing interests exist.

Ethical statement

No animal was captured, handled or killed specifically for the purpose of this study. Reference samples of *Ovis aries* originated from a meat processing plant, while *Ovis musimon* samples were obtained from legally harvested individuals during authorised hunting activities. Neither species is protected in France and, therefore, no specific permits were required for their collection.

Wolf (*Canis lupus*) scat samples were collected non-invasively in the field as part of the national wolf monitoring programme coordinated by the French Agency for Biodiversity, under the authority of the French government. Scat collection did not involve any disturbance, capture or handling of animals. All sampling and analyses complied with relevant national regulations and ethical guidelines.

Use of AI

No use of AI was reported.

Funding

This study was partly funded by the Research and Development Contract N° OFB-23-1135 of the French Agency for Biodiversity and by the Public Contract for the Supply of Genetic Analysis for the Monitoring of Wolf (*Canis lupus*) Populations N° 2021-28.

Author contributions

Conceptualization: CD, GQ. Data curation: GQ, AP, SR, JL, CK. Formal analysis: AP, GQ, CK, JL, SR. Funding acquisition: GQ, CD. Project administration: CD, GQ. Supervision: CD, GQ. Visualization: AP. Writing - original draft: AP. Writing - review and editing: AP, JL, SR, GQ, CD, CK.

Author ORCIDs

Agathe Pirog  <https://orcid.org/0000-0001-7994-1134>

Cécile Kaerle  <https://orcid.org/0000-0002-4078-0553>

Christophe Duchamp  <https://orcid.org/0000-0002-0448-3412>

Guillaume Queney  <https://orcid.org/0000-0002-6071-5169>

Data availability

Whole D-loop haplotypes (674 bp) from reference *Ovis* samples have been deposited in GenBank under accession numbers [PX219566](#) to [PX219610](#) and raw Ion Torrent

sequencing data have been deposited in the European Nucleotide Archive (ENA) under project accession PRJEB106338.

All supplementary materials as well as OTUs/ASVs tables and the *Ovis* reference database (OvisCR_ref) are available on Zenodo (DOI: 10.5281/zenodo.18185713).

References

- Alberdi A, Aizpurua O, Gilbert MTP, Bohmann K (2018) Scrutinizing key steps for reliable metabarcoding of environmental samples. *Methods in Ecology and Evolution* 9: 134–147. <https://doi.org/10.1111/2041-210X.12849>
- Ando H, Mukai H, Komura T, Dewi T, Ando M, Isagi Y (2020) Methodological trends and perspectives of animal dietary studies by noninvasive fecal DNA metabarcoding. *Environmental DNA* 2: 391–406. <https://doi.org/10.1002/edn3.117>
- Barbato M, Hailer F, Orozco-terWengel P, Kijas J, Mereu P, Cabras P, Mazza R, Pirastru M, Bruford MW (2017) Genomic signatures of adaptive introgression from European mouflon into domestic sheep. *Scientific Reports* 7: 7623. <https://doi.org/10.1038/s41598-017-07382-7>
- Bohmann K, Elbrecht V, Carøe C, Bista I, Leese F, Bunce M, Yu DW, Seymour M, Dumbrell AJ, Creer S (2022) Strategies for sample labelling and library preparation in DNA metabarcoding studies. *Molecular Ecology Resources* 22: 1231–1246. <https://doi.org/10.1111/1755-0998.13512>
- Boyer F, Mercier C, Bonin A, Le Bras Y, Taberlet P, Coissac E (2016) OBITools: A unix-inspired software package for DNA metabarcoding. *Molecular Ecology Resources* 16: 176–182. <https://doi.org/10.1111/1755-0998.12428>
- Bruns A, Waltert M, Khorozyan I (2020) The effectiveness of livestock protection measures against wolves (*Canis lupus*) and implications for their co-existence with humans. *Global Ecology and Conservation* 21: e00868. <https://doi.org/10.1016/j.gecco.2019.e00868>
- Buzan E, Pokorný B, Urzi F, Duniš L, Bončina A, Iacolina L, Šprem N, Stipoljev S, Mereu P, Leoni G, Pirastru M, Safner T (2024a) Genetic variation of European mouflon depends on admixture of introduced individuals. *Mammal Research* 69: 145–158. <https://doi.org/10.1007/s13364-023-00726-x>
- Buzan E, Potočník H, Pokorný B, Potušek S, Iacolina L, Gerič U, Urzi F, Kos I (2024b) Molecular analysis of scats revealed diet and prey choice of grey wolves and Eurasian lynx in the contact zone between the Dinaric Mountains and the Alps. *Frontiers in Zoology* 21: 9. <https://doi.org/10.1186/s12983-024-00530-6>
- Callahan BJ, McMurdie PJ, Rosen MJ, Han AW, Johnson AJA, Holmes SP (2016) DADA2: High-resolution sample inference from Illumina amplicon data. *Nature Methods* 13: 581–583. <https://doi.org/10.1038/nmeth.3869>
- Callahan BJ, McMurdie PJ, Holmes SP (2017) Exact sequence variants should replace operational taxonomic units in marker-gene data analysis. *The ISME Journal* 11: 2639–2643. <https://doi.org/10.1038/ismej.2017.119>
- Casper RM, Jarman SN, Deagle BE, Gales NJ, Hindell MA (2007) Detecting prey from DNA in predator scats: A comparison with morphological analysis, using *Arctocephalus* seals fed a known diet. *Journal of Experimental Marine Biology and Ecology* 347: 144–154. <https://doi.org/10.1016/j.jembe.2007.04.002>
- Chapron G, Kaczensky P, Linnell JDC, von Arx M, Huber D, Andrén H, López-Bao JV, Adamec M, Álvares F, Anders O, Balčiauskas L, Balys V, Bedő P, Bego F, Blanco JC, Breitenmoser U, Brøseth H, Bufka L, Bunikyte R, Ciucci P, Dutsov A, Engleder T, Fuxjäger

- C, Groff C, Holmala K, Hoxha B, Iliopoulos Y, Ionescu O, Jeremić J, Jerina K, Kluth G, Knauer F, Kojola I, Kos I, Krofel M, Kubala J, Kunovac S, Kusak J, Kutal M, Liberg O, Majić A, Männil P, Manz R, Marboutin E, Marucco F, Melovski D, Mersini K, Mertzanis Y, Mysłajek RW, Nowak S, Odden J, Ozolins J, Palomero G, Paunović M, Persson J, Potočnik H, Quenette P-Y, Rauer G, Reinhardt I, Rigg R, Ryser A, Salvatori V, Skrbineš T, Stojanov A, Swenson JE, Szemethy L, Trajçe A, Tsingarska-Sedefcheva E, Váňa M, Veeroja R, Wabakken P, Wölfl M, Wölfl S, Zimmermann F, Zlatanova D, Boitani L (2014) Recovery of large carnivores in Europe's modern human-dominated landscapes. *Science* 346: 1517–1519. <https://doi.org/10.1126/science.1257553>
- Chen Z-H, Xu Y-X, Xie X-L, Wang D-F, Aguilar-Gómez D, Liu G-J, Li X, Esmailizadeh A, Rezaei V, Kantanen J, Ammosov I, Nosrati M, Periasamy K, Coltman DW, Lenstra JA, Nielsen R, Li M-H (2021) Whole-genome sequence analysis unveils different origins of European and Asiatic mouflon and domestication-related genes in sheep. *Communications Biology* 4: 1307. <https://doi.org/10.1038/s42003-021-02817-4>
- Ciani E, Mastrangelo S, Da Silva A, Marroni F, Ferenčaković M, Ajmone-Marsan P, Baird H, Barbato M, Colli L, Delvento C, Dovenski T, Gorjanc G, Hall SJG, Hoda A, Li M-H, Marković B, McEwan J, Moradi MH, Ruiz-Larrañaga O, Ružić-Muslić D, Šalamon D, Simčič M, Stepanek O, Curik I, Cubric-Curik V, Lenstra JA, Econogene C, Sheephapmap C (2020) On the origin of European sheep as revealed by the diversity of the Balkan breeds and by optimizing population-genetic analysis tools. *Genetics, Selection, Evolution : GSE* 52: 25. <https://doi.org/10.1186/s12711-020-00545-7>
- Ciucci P, Boitani L, Pelliccioni ER, Rocco M, Guy I (1996) A comparison of scat-analysis methods to assess the diet of the wolf *Canis lupus*. *Wildlife Biology* 2: 37–48. <https://doi.org/10.2981/wlb.1996.006>
- Clement M, Posada D, Crandall KA (2000) TCS: A computer program to estimate gene genealogies. *Molecular Ecology* 9: 1657–1659. <https://doi.org/10.1046/j.1365-294x.2000.01020.x>
- Couton M, Baud A, Daguin-Thiébaud C, Corre E, Comtet T, Viard F (2021) High-throughput sequencing on preservative ethanol is effective at jointly examining infraspecific and taxonomic diversity, although bioinformatics pipelines do not perform equally. *Ecology and Evolution* 11: 5533–5546. <https://doi.org/10.1002/ece3.7453>
- de Sousa LL, Silva SM, Xavier R (2019) DNA metabarcoding in diet studies: Unveiling ecological aspects in aquatic and terrestrial ecosystems. *Environmental DNA* 1: 199–214. <https://doi.org/10.1002/edn3.27>
- Deagle BE, Thomas AC, Shaffer AK, Trites AW, Jarman SN (2013) Quantifying sequence proportions in a DNA-based diet study using Ion Torrent amplicon sequencing: Which counts count? *Molecular Ecology Resources* 13: 620–633. <https://doi.org/10.1111/1755-0998.12103>
- Di Bernardi C, Wikenros C, Hedmark E, Boitani L, Ciucci P, Sand H, Åkesson M (2021) Multiple species-specific molecular markers using nanofluidic array as a tool to detect prey DNA from carnivore scats. *Ecology and Evolution* 11: 11739–11748. <https://doi.org/10.1002/ece3.7918>
- Drake LE, Cuff JP, Young RE, Marchbank A, Chadwick EA, Symondson WOC (2022) An assessment of minimum sequence copy thresholds for identifying and reducing the prevalence of artefacts in dietary metabarcoding data. *Methods in Ecology and Evolution* 13: 694–710. <https://doi.org/10.1111/2041-210X.13780>
- Duchamp C, Boyer J, Briaudet P-E, Leonard Y, Moris P, Bataille A, Dahier T, Delacour G, Millisher G, Miquel C, Poillot C, Marboutin E (2012) Wolf monitoring in France: A dual frame process to survey time- and space-related changes in the population.

- Hystrix, the Italian Journal of Mammalogy 23: 14–28. <https://doi.org/10.4404/hystrix-23.1-4559>
- Feng J, Frisina MR, Webster MS, Ulziimaa G (2009) Genetic differentiation of argali sheep *Ovis ammon* in Mongolia revealed by mitochondrial control region and nuclear microsatellites analyses. Journal of the Bombay Natural History Society 106: 38–44.
- Forin-Wiart M-A, Poulle M-L, Piry S, Cosson J-F, Larose C, Galan M (2018) Evaluating metabarcoding to analyse diet composition of species foraging in anthropogenic landscapes using Ion Torrent and Illumina sequencing. Scientific Reports 8: 17091. <https://doi.org/10.1038/s41598-018-34430-7>
- Gervasi V, Linnell JDC, Berce T, Boitani L, Cerne R, Ciucci P, Cretois B, Derron-Hilfiker D, Duchamp C, Gastineau A, Grente O, Huber D, Iliopoulos Y, Karamanlidis AA, Kojola I, Marucco F, Mertzanis Y, Männil P, Norberg H, Pagon N, Pedrotti L, Quenette P-Y, Reljic S, Salvatori V, Talvi T, von Arx M, Gimenez O (2021) Ecological correlates of large carnivore depredation on sheep in Europe. Global Ecology and Conservation 30: e01798. <https://doi.org/10.1016/j.gecco.2021.e01798>
- Gouy M, Guindon S, Gascuel O (2009) Seaview version 4: A multiplatform graphical user interface for sequence alignment and phylogenetic tree building. Molecular Biology and Evolution 27: 221–224. <https://doi.org/10.1093/molbev/msp259>
- Grente O, Saubusse T, Gimenez O, Marboutin E, Duchamp C (2022) Wolf depredation hotspots in France: Clustering analyses adjusting for livestock availability. Biological Conservation 267: 109495. <https://doi.org/10.1016/j.biocon.2022.109495>
- Hacker CE, Jevit M, Hussain S, Muhammad G, Munkhtsog B, Munkhtsog B, Zhang Y, Li D, Liu Y, Farrington JD, Balbakova F, Alamanov A, Kurmanaliev O, Buyanaa C, Bayandonoi G, Ochirjav M, Liang X, Bian X, Weckworth B, Jackson R, Janecka JE (2021) Regional comparison of snow leopard (*Panthera uncia*) diet using DNA metabarcoding. Biodiversity and Conservation 30: 797–817. <https://doi.org/10.1007/s10531-021-02118-6>
- Hakimzadeh A, Abdala Asbun A, Albanese D, Bernard M, Buchner D, Callahan B, Caporaso JG, Curd E, Djemiel C, Brandström Durling M, Elbrecht V, Gold Z, Gweon HS, Hajibabaei M, Hildebrand F, Mikryukov V, Normandeau E, Özkurt E, M Palmer J, Pascal G, Porter TM, Straub D, Vasar M, Větrovský T, Zafeiropoulos H, Anslan S (2024) A pile of pipelines: An overview of the bioinformatics software for metabarcoding data analyses. Molecular Ecology Resources 24: e13847. <https://doi.org/10.1111/1755-0998.13847>
- Hiendleder S, Mainz K, Plante Y, Lewalski H (1998) Analysis of mitochondrial DNA indicates that domestic sheep are derived from two different ancestral maternal sources: No evidence for contributions from urial and argali sheep. The Journal of Heredity 89: 113–120. <https://doi.org/10.1093/jhered/89.2.113>
- Hiendleder S, Kaupe B, Wassmuth R, Janke A (2002) Molecular analysis of wild and domestic sheep questions current nomenclature and provides evidence for domestication from two different subspecies. Proceedings of the Royal Society of London. Series B, Biological Sciences 269: 893–904. <https://doi.org/10.1098/rspb.2002.1975>
- IDELE (2018) Ovins 2017 Productions lait et viande, Les chiffres clés du GEB.
- Jusino MA, Banik MT, Palmer JM, Wray AK, Xiao L, Pelton E, Barber JR, Kawahara AY, Gratton C, Peery MZ, Lindner DL (2019) An improved method for utilizing high-throughput amplicon sequencing to determine the diets of insectivorous animals. Molecular Ecology Resources 19: 176–190. <https://doi.org/10.1111/1755-0998.12951>
- Kearse M, Moir R, Wilson A, Stones-Havas S, Cheung M, Sturrock S, Buxton S, Cooper A, Markowitz S, Duran C, Thierer T, Ashton B, Meintjes P, Drummond A (2012) Geneious Basic: An integrated and extendable desktop software platform for the organization

- and analysis of sequence data. *Bioinformatics* (Oxford, England) 28: 1647–1649. <https://doi.org/10.1093/bioinformatics/bts199>
- Khalatbari L, Egeter B, Abolghasemi H, Hakimi E, Ghadirian T, Khaleghi Hamidi AH, Jowkar H, Breitenmoser U, Brito JC (2022) Assessing Asiatic cheetah’s individual diet using metabarcoding and its implication for conservation. *Scientific Reports* 12: 11403. <https://doi.org/10.1038/s41598-022-15065-1>
- Klymus KE, Baker JD, Abbott CL, Brown RJ, Craine JM, Gold Z, Hunter ME, Johnson MD, Jones DN, Jungbluth MJ, Lor Y, Maloy AP, Merkes CM, Noble R, Patin NV, Sepulveda AJ, Spear SF, Steele JA, Takahashi M, Watts AW, Theroux S (2024) The MIEM guidelines: Minimum information for reporting of environmental metabarcoding data. *Metabarcoding and Metagenomics* 8. <https://doi.org/10.3897/mbmg.8.128689>
- Lawson Handley LJ, Byrne K, Santucci F, Townsend S, Taylor M, Bruford MW, Hewitt GM (2007) Genetic structure of European sheep breeds. *Heredity* 99: 620–631. <https://doi.org/10.1038/sj.hdy.6801039>
- Leigh JW, Bryant D, Nakagawa S (2015) PopART: Full-feature software for haplotype network construction. *Methods in Ecology and Evolution* 6(9): 1110–1116. <https://doi.org/10.1111/2041-210X.12410>
- Leopold BD, Krausman PR (1986) Diets of 3 Predators in Big Bend National Park, Texas. *The Journal of Wildlife Management* 50: 290–295. <https://doi.org/10.2307/3801915>
- Linnell J, Cretois B (2018) Research for AGRI Committee - The revival of wolves and other large predators and its impact on farmers and their livelihood in rural regions of Europe. European Parliament, Brussels.
- Littleford-Colquhoun BL, Freeman PT, Sackett VI, Tulloss CV, McGarvey LM, Geremia C, Kartzinel TR (2022) The precautionary principle and dietary DNA metabarcoding: Commonly used abundance thresholds change ecological interpretation. *Molecular Ecology* 31: 1615–1626. <https://doi.org/10.1111/mec.16352>
- Massey AL, Roffler GH, Vermeul T, Allen JM, Levi T (2021) Comparison of mechanical sorting and DNA metabarcoding for diet analysis with fresh and degraded wolf scats. *Ecosphere* 12: e03557. <https://doi.org/10.1002/ecs2.3557>
- Meadows JRS, Hiendleder S, Kijas JW (2011) Haplogroup relationships between domestic and wild sheep resolved using a mitogenome panel. *Heredity* 106: 700–706. <https://doi.org/10.1038/hdy.2010.122>
- Morehouse AT, Boyce MS (2017) Troublemaking carnivores: Conflicts with humans in a diverse assemblage of large carnivores. *Ecology and Society* 22(3): 4. <https://doi.org/10.5751/ES-09415-220304>
- Nadler CF, Hoffmann RS, Woolf A (1973) G-band patterns as chromosomal markers, and the interpretation of chromosomal evolution in wild sheep (*Ovis*). *Experientia* 29: 117–119. <https://doi.org/10.1007/BF01913288>
- Paradis E (2010) pegas: An R package for population genetics with an integrated–modular approach. *Bioinformatics* (Oxford, England) 26: 419–420. <https://doi.org/10.1093/bioinformatics/btp696>
- Pirog A, Duchamp C, Kaerle C, Dufaure de Citres C, Rousselot S, Lavarec J, Queney G (2025) Standardization of a high-quality methodological framework for long-term genetic monitoring of the French wolf population. *Ecology and Evolution* 15: e71345. <https://doi.org/10.1002/ece3.71345>
- Portanier E, Chevret P, Gélín P, Benedetti P, Sanchis F, Barbanera F, Kaerle C, Queney G, Bourgoin G, Devillard S, Garel M (2022) New insights into the past and recent evolutionary history of the Corsican mouflon (*Ovis gmelini musimon*) to inform its conservation. *Conservation Genetics* 23: 91–107. <https://doi.org/10.1007/s10592-021-01399-2>

- R Core Team (2023) R: A language and environment for statistical computing. R Foundation for Statistical Computing, Vienna, Austria.
- Reynolds JC, Aebischer NJ (1991) Comparison and quantification of carnivore diet by faecal analysis: A critique, with recommendations, based on a study of the Fox *Vulpes vulpes*. *Mammal Review* 21: 97–122. <https://doi.org/10.1111/j.1365-2907.1991.tb00113.x>
- Rezaei HR, Naderi S, Chintauan-Marquier IC, Taberlet P, Virk AT, Naghash HR, Rioux D, Kaboli M, Pompanon F (2010) Evolution and taxonomy of the wild species of the genus *Ovis* (Mammalia, Artiodactyla, Bovidae). *Molecular Phylogenetics and Evolution* 54: 315–326. <https://doi.org/10.1016/j.ympev.2009.10.037>
- Riaz T, Shehzad W, Viari A, Pompanon F, Taberlet P, Coissac E (2011) ecoPrimers: Inference of new DNA barcode markers from whole genome sequence analysis. *Nucleic Acids Research* 39: e145–e145. <https://doi.org/10.1093/nar/gkr732>
- Ripple WJ, Estes JA, Beschta RL, Wilmers CC, Ritchie EG, Hebblewhite M, Berger J, Elmhagen B, Letnic M, Nelson MP, Schmitz OJ, Smith DW, Wallach AD, Wirsing AJ (2014) Status and ecological effects of the world's largest carnivores. *Science* 343: 1241484. <https://doi.org/10.1126/science.1241484>
- Rochus CM, Tortereau F, Plisson-Petit F, Restoux G, Moreno-Romieux C, Tosser-Klopp G, Servin B (2018) Revealing the selection history of adaptive loci using genome-wide scans for selection: An example from domestic sheep. *BMC Genomics* 19: 71. <https://doi.org/10.1186/s12864-018-4447-x>
- Roffler GH, Allen JM, Massey A, Levi T (2021) Metabarcoding of fecal DNA shows dietary diversification in wolves substitutes for ungulates in an island archipelago. *Ecosphere* 12: e03297. <https://doi.org/10.1002/ecs2.3297>
- Royaux C, Norvez O, Coissac E, Boyer F, Le Bras Y (2022) Metabarcoding/eDNA through Obitools (Galaxy Training Materials). <https://training.galaxyproject.org/training-material/topics/ecology/tutorials/Obitools-metabarcoding/tutorial.html>
- Ryder M (1991) Domestication, history and breed evolution in sheep. Volume B8. Elsevier, Amsterdam.
- Savouré-Soubelet A, Arthur C, Aulagnier S, Body G, Callou C, Haffner P, Marchandeu S, Moutou F, Saint-Andrieux C (2021) Atlas des mammifères sauvages de France volume 2: Ongulés et Lagomorphes. Volume 83. Muséum national d'Histoire naturelle, Paris.
- Schroeder H, Palczewski S, Degen B (2021) Development of D-Loop mitochondrial markers for amplification of prey DNA from wolf scat. *Conservation Genetics Resources* 13: 1–4. <https://doi.org/10.1007/s12686-020-01169-1>
- Shi Y, Hoareau Y, Reese EM, Wasser SK (2021) Prey partitioning between sympatric wild carnivores revealed by DNA metabarcoding: A case study on wolf (*Canis lupus*) and coyote (*Canis latrans*) in northeastern Washington. *Conservation Genetics* 22: 293–305. <https://doi.org/10.1007/s10592-021-01337-2>
- Simon C, Buckley TR, Frati F, Stewart JB, Beckenbach AT (2006) Incorporating molecular evolution into phylogenetic analysis, and a new compilation of conserved polymerase chain reaction primers for animal mitochondrial DNA. *Annual Review of Ecology, Evolution, and Systematics* 37: 545–579. <https://doi.org/10.1146/annurev.ecolsys.37.091305.110018>
- Spaulding R, Krausman PR, Ballard WB (2000) Observer bias and analysis of gray wolf diets from scats. *Wildlife Society Bulletin* 28: 947–950.

Terrel MPTG, Cuff JP (2022) The complex epistemological challenge of data curation in dietary metabarcoding: Comment on “The precautionary principle and dietary DNA metabarcoding: Commonly used abundance thresholds change ecological interpretation” by Littleford-Colquhoun et al. (2022). *Molecular Ecology* 31: 5653–5659. <https://doi.org/10.1111/mec.16576>

Yang L, Tan Z, Wang D, Xue L, Guan M-x, Huang T, Li R (2014) Species identification through mitochondrial rRNA genetic analysis. *Scientific Reports* 4: 4089. <https://doi.org/10.1038/srep04089>

Supplementary material 1

DNA quantification protocols

Authors: Agathe Pirog, Juliette Lavarec, Sabine Rousselot, Cécile Kaerle, Christophe Duchamp, Guillaume Queney

Data type: pdf

Explanation note: Protocols used to quantify DNA for metabarcoding libraries protocol.

Copyright notice: This dataset is made available under the Open Database License (<http://opendatacommons.org/licenses/odbl/1.0/>). The Open Database License (ODbL) is a license agreement intended to allow users to freely share, modify, and use this Dataset while maintaining this same freedom for others, provided that the original source and author(s) are credited.

Link: <https://doi.org/10.3897/mbmg.10.172476.suppl1>

Supplementary material 2

Functions and parameters used in the metabarcoding pipelines

Authors: Agathe Pirog, Juliette Lavarec, Sabine Rousselot, Cécile Kaerle, Christophe Duchamp, Guillaume Queney

Data type: xlsx

Explanation note: Functions and main parameters used for wolf scats metabarcoding analyses.

Copyright notice: This dataset is made available under the Open Database License (<http://opendatacommons.org/licenses/odbl/1.0/>). The Open Database License (ODbL) is a license agreement intended to allow users to freely share, modify, and use this Dataset while maintaining this same freedom for others, provided that the original source and author(s) are credited.

Link: <https://doi.org/10.3897/mbmg.10.172476.suppl2>

Supplementary material 3

Diagnostic Dloop fragment (OvisCR) haplotypes (178pb) for 92 *Ovis* reference samples collected in France

Authors: Agathe Pirog, Juliette Lavarec, Sabine Rousselot, Cécile Kaerle, Christophe Duchamp, Guillaume Queney

Data type: xlsx

Explanation note: Information on 92 *Ovis* reference samples, including:- population of origin for European mouflon samples- breed for domestic sheep samples- name of the complete control region haplotype (674 bp)- name and sequence of the corresponding OvisCR fragment haplotype (178 bp).

Copyright notice: This dataset is made available under the Open Database License (<http://opendatacommons.org/licenses/odbl/1.0/>). The Open Database License (ODbL) is a license agreement intended to allow users to freely share, modify, and use this Dataset while maintaining this same freedom for others, provided that the original source and author(s) are credited.

Link: <https://doi.org/10.3897/mbmg.10.172476.suppl3>

Supplementary material 4

TCS haplotype network based on a 674 bp D-loop fragment, constructed from 30 *Ovis musimon* and 41 *Ovis aries* reference samples collected in France

Authors: Agathe Pirog, Juliette Lavarec, Sabine Rousselot, Cécile Kaerle, Christophe Duchamp, Guillaume Queney

Data type: pdf

Explanation note: TCS haplotype network, based on a 674 base pair diagnostic D-loop fragment, constructed from 30 *Ovis musimon* and 41 *Ovis aries* reference samples collected in France. Each haplotype is represented by a circle for *O. aries* and a circle enclosed within a square for *O. musimon*. Each tick mark on the lines connecting the haplotypes indicates a single mutation. The circle size is proportional to the number of individuals sharing that haplotype and the colours indicate the species

Copyright notice: This dataset is made available under the Open Database License (<http://opendatacommons.org/licenses/odbl/1.0/>). The Open Database License (ODbL) is a license agreement intended to allow users to freely share, modify, and use this Dataset while maintaining this same freedom for others, provided that the original source and author(s) are credited.

Link: <https://doi.org/10.3897/mbmg.10.172476.suppl4>

Supplementary material 5

Metabarcoding results of 115 wolf scat samples collected in France and sequenced using the 12S and OvisCR fragments

Authors: Agathe Pirog, Juliette Lavarec, Sabine Rousselot, Cécile Kaerle, Christophe Duchamp, Guillaume Queney

Data type: xlsx

Explanation note: Diet metabarcoding results from 115 wolf (*Canis lupus*) scat samples collected in France, sequenced with both the universal 12S fragment and the diagnostic OvisCR fragment. Metadata and results include:- French department (number) of sampling,- sex of the wolf (M or F),- total number of reads and *Ovis* reads obtained with the 12S fragment,- number of reads obtained with the OvisCR fragment,- species-level assignments of *Ovis* sequences. Missing data are indicated "NA".

Copyright notice: This dataset is made available under the Open Database License (<http://opendatacommons.org/licenses/odbl/1.0/>). The Open Database License (ODbL) is a license agreement intended to allow users to freely share, modify, and use this Dataset while maintaining this same freedom for others, provided that the original source and author(s) are credited.

Link: <https://doi.org/10.3897/mbmg.10.172476.suppl5>